④

TECHNICAL REPORT

Office of Naval Research Contract No. N00014-86-K0029

ROOTFINDING FOR MARKOV CHAINS WITH QUASI-TRIANGULAR TRANSITION MATRICES

by

Carl M. Harris

Report No. GMU/22472/105
October 31, 1988

Department of Operations Research and Applied Statistics
School of Information Technology and Engineering
George Mason University
Fairfax, VA 22030

TECHNICAL REPORT

Office of Naval Research Contract No. N00014-86-K0029

ROOTFINDING FOR MARKOV CHAINS WITH QUASI-TRIANGULAR TRANSITION MATRICES

by

Carl M. Harris

Report No. GMU/22472/105
October 31, 1988

Department of Operations Research and Applied Statistics
School of Information Technology and Engineering
George Mason University
Fairfax, VA 22030

Copy No. 13

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE

READ INSTRUCTIONS BEFORE COMPLETING FORM

1. REPORT NUMBER	2. GOVT ACCESSION NO.	3. RECIPIENT'S CATALOG NUMBER
4. TITLE (and Subtitle) Rootfinding for Markov Chains with Quasi-Triangular Transition Matrices		5. TYPE OF REPORT & PERIOD COVERED Technical Report
		6. PERFORMING ORG. REPORT NUMBER GMU/22474/105
7. AUTHOR(s) Carl M. Harris		8. CONTRACT OR GRANT NUMBER(s) N00014-86-K0029
9. PERFORMING ORGANIZATION NAME AND ADDRESS Department of Operations Research and Applied Statistics George Mason University, Fairfax, Va. 22030		10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS Task H-B Project 4118150
11. CONTROLLING OFFICE NAME AND ADDRESS Office of Naval Research 800 North Quincy Street Arlington, VA. 22247		12. REPORT DATE October 31, 1988
		13. NUMBER OF PAGES 9
14. MONITORING AGENCY NAME & ADDRESS (if different from Controlling Office)		15. SECURITY CLASS. (of this report) Unclassified
		15a. DECLASSIFICATION/DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT (of this Report)

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

18. SUPPLEMENTARY NOTES

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

applied probability, computational analysis, computational probability Markov chains numerical methods, probability, rootfinding stochastic models,

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

(following)

DD FORM 1 JAN 73 1473 EDITION OF 1 NOV 65 IS OBSOLETE

S/N 0102-LF-014-6601

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

Abstract

Numerical rootfinding problems are quite common in stochastic modeling. However, many solutions stop at the presentation of a probability generating function for the state probabilities. But with increasing easy access to computing power, many problems whose answers were typically left in incomplete form or for which there has been a search for alternative solution methods are currently being reexamined. The class of Markov chains with quasi-triangular layouts (i.e., those having sub- or super-triangular sets of zeros) are a good case in point. They have an especially nice structure which leads to a rather concise representation for the generating functions. But the complete solution then requires the finding of roots. Fortunately, these problems can be shown to have special properties that make accurate rootfinding quite feasible. In this paper, we show that the roots of the critical equations for these models are indeed unique and located in known regions in the complex plane.

1 INTRODUCTION

As Abolnikov and Dukhovny (1987) and others (for example, see Bailey, 1954, and Powell, 1985) have repeatedly noted, many denumerable discrete-time Markov chains (with particular applications in inventory, dam and queueing modeling) have one of two special transition-matrix structures. These forms have been called *quasi-triangular* by a number of authors, because of the presence of sub- or super-triangular sets of zeros:

$$A = \begin{bmatrix} a_{00} & a_{01} & a_{02} & a_{03} & \cdots \\ a_{10} & a_{11} & a_{12} & a_{13} & \cdots \\ a_{20} & a_{21} & a_{22} & a_{23} & \cdots \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ a_{K0} & a_{K1} & a_{K2} & a_{K3} & \cdots \\ a_0 & a_1 & a_2 & a_3 & \cdots \\ 0 & a_0 & a_1 & a_2 & \cdots \\ 0 & 0 & a_0 & a_1 & \cdots \\ 0 & 0 & 0 & a_0 & \cdots \\ 0 & 0 & 0 & 0 & \cdots \\ \cdots & \cdots & \cdots & \cdots & \cdots \end{bmatrix}$$

and

$$B = \begin{bmatrix} \Sigma_{K-1} & b_{K-1} & b_{K-2} & b_{K-3} & \cdots & b_0 & 0 & 0 & 0 & \cdots \\ \Sigma_K & b_K & b_{K-1} & b_{K-2} & \cdots & b_1 & b_0 & 0 & 0 & \cdots \\ \Sigma_{K+1} & b_{K+1} & b_K & b_{K-1} & \cdots & b_2 & b_1 & b_0 & 0 & \cdots \\ \Sigma_{K+2} & b_{K+2} & b_{K+1} & b_K & \cdots & b_3 & b_2 & b_1 & b_0 & \cdots \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \end{bmatrix},$$

where

$$\Sigma_j = 1 - \sum_{n=0}^{j} b_n.$$

The structure of these matrices leads to some particularly concise representations for the probability generating functions (PGFs) of the Markov chain equilibrium state process. When the stationary equation for such a

Markov chain is exercised, the PGF of the steady-state probabilities has an algebraic function in its denominator whose roots are critical in the final solution. (We henceforth refer to this denominator function equated to zero as the system's *characteristic equation*).

For the transition matrix **A**, the characteristic equation (CE) turns out to be (at least for complex z with absolute value ≤ 1)

$$z^K = \sum_{i=0}^{\infty} a_i z^i = \alpha(z), \tag{1}$$

where K is as defined in the matrix representation **A** (corresponding, for example, to a constant batch-input module in bulk queues). Under the assumption that $\alpha(z)$ possesses all its derivatives at $z = 1$ (i.e., that the distribution $\{a_i\}$ has all moments), $\alpha(z)$ may be set equal to the Laplace-Stieltjes transform of a distribution function [call that A(t), and set its mean to $1/\mu$] of a nonnegative random variable evaluated at $\lambda(1 - z)$, where λ is an arbitrary positive constant for the time being. Thus we may also write that

$$z^K = A^*[\lambda(1 - z)]. \tag{2}$$

The CE associated with the matrix **B** may be written as

$$z^K = \sum_{i=0}^{\infty} b_i z^i = \beta(z), \tag{3}$$

or

$$z^K = B^*[\mu(1 - z)]. \tag{4}$$

In these representations, the constant K is as given in **B** (corresponding, for example, to a constant batch-service module in bulk queues), $\beta(z)$ is defined as the PGF of the probabilities $\{b_i\}$, and B^* is the Laplace-Stieltjes transform of a distribution function [call that B(t), and set its mean to $1/\lambda$] of a nonnegative random variable evaluated at $\mu(1 - z)$, where μ is an arbitrary positive constant.

Recognize that Equations (1) and (3) are generalizations of the well-known fundamental equation of branching processes, typically written as $z = f(z)$, where f would be the PGF for the number of offspring emanating from one parent. Gross and Harris (1985), for example, provide the details of the root problem for this model. In actuality, it is the **B** problem which is in fact the more direct relative of the branching process, so it is this one on which we comment in detail first.

The CE of the matrix **B** may be rewritten in the standard way by using $z = r exp(i\theta)$, and we find that

$$r^K exp(i\theta K) = B^*[\mu(1 - r\, exp(i\theta))] exp(2\pi n i) \tag{5}$$

for $n = 1, 2, ..., K$. This equation clearly has a root at unity, and by Rouché's theorem, we can show that there are K others inside the unit circle $|z| = 1$ when the chain is ergodic. The condition for ergodicity is that

$$\beta'(1) = \sum_{n=0}^{\infty} n b_n > K$$

or

$$\left(\frac{dB^*[\mu(1-z))]}{dz}\right)_{z=1} > K.$$

These can be shown to be equivalent to the requirement that $K\lambda/\mu < 1$. In addition and most importantly, for each n in (5), there is a unique root with absolute value less than 1. We provide the proof of this assertion in the next section.

The roots of (5) are found by separately solving its real and imaginary portions. As noted, there is a unique answer when (5) is evaluated for individual values of n. In Section 2, we use this fact to complete the proof that these roots are indeed unique. In addition, we are able to pinpoint the locations of all of the real roots.

Partial results on the uniqueness of all the roots for this model type are available. It is, in fact, known that the roots are totally unique when $B(t)$ is Erlang(J) distributed with mean $1/\lambda$ (see Chaudhry and Templeton, 1983). In this case, Equation 4 becomes (for $K\lambda/\mu = K\rho < 1$)

$$z^K = [J\rho/(J\rho + (1 - z)]^J. \tag{6}$$

To prove that the roots are simple, we show that the derivative vanishes at a potential repeated root, say z_i. This is equivalent to requiring that

$$K z_i^{K-1} = J(J\rho)^J [J\rho + (1 - z_i)]^{-(J+1)}. \tag{7}$$

If (6) is divided by (7), we see that

$$z_i = \frac{K(J\rho+1)}{J+K} = 1 - \frac{J(1-K\rho)}{J+K}.$$

When this value is substituted back into Equation (6), we find that the condition for repetition is that

$$(K\rho)^J = (J\rho+1)^{J+K}(\frac{K}{J+K})^{J+K}$$

or

$$(K\rho)^{J/(J+K)} = (J\rho+1)(\frac{K}{J+K}).$$

If we define $J/(J+K)$ as c, then the condition for repetition is equivalent to requiring that

$$(K\rho)^c = (J\rho+1)(1-c) = (1-c)J\rho + (1-c). \tag{8}$$

The right-hand side of (8) is a straight line in ρ, with y-intercept of 1-c and slope of (1-c)J, while the left-hand side is a monomial with positive, fractional power c. The two functions intersect only at $\rho = 1/K$, which would violate the condition for ergodicity that $\rho < 1/K$, so that the assumption of repetition must be false.

But not much more has been generally known about the effect of the form of the distribution $B(t)$ for the more general model.

For the A-matrix problem, recall that the characteristic equation is

$$z^K = A^*[\lambda(1-z)], \tag{9}$$

where A^* is a Laplace-Stieltjes transform. Ergodocity obtains here when

$$\alpha'(1) = \sum_{n=0}^{\infty} na_n < K$$

or

$$\left(\frac{dA^*[\lambda(1-z)]}{dz}\right)_{z=1} < K.$$

This is equivalent to requiring that $\lambda/K\mu < 1$.

It is easily shown by Rouché's theorem that (9) has K roots inside and on the unit circle, including the root z=1. Abolnikov and Dukhovny (1987) have noted that all the roots on the unit circle are, in fact, simple. We show here that the specific root z=1 is simple using the usual derivative test. To do so, we evaluate

$$Kz^{K-1} = -\lambda dA^*(\lambda - \lambda z)/dz$$

at z=1 and find that

$$K = -\lambda(-1/\mu) = \rho,$$

or $\rho = K$, which is a contradiction of the ergodic condition that $\rho < K$. Hence z=1 cannot be a double root.

When $A(t)$ is deterministic, we can further show that all K-1 roots inside or on the unit circle are, in fact, strictly within. This follows when we rewrite (13) as

$$z^K = e^{K\rho(z-1)}$$

and assume that z_i has absolute value of 1 but is not precisely equal to 1. Then we see that

$$1 = e^{K\rho(z_i-1)},$$

which implies that $Re(z_i - 1) = 0$ and thus that $z_i = 1$. But this is contrary to our earlier verification that the root $z = 1$ is simple.

2 NEW RESULTS

We first show:

Lemma 1 *Equation 5 has one and only one root for each n.*

Proof: Rewrite (5) as

$$r\,exp[i(\theta - \frac{2\pi n}{K})] = \left[\int_0^\infty e^{-\mu[1-r\,exp(i\theta)t]}\,dB(t)\right]^{1/K}.$$

Let $f(z)$ equal the left-hand side of the above equation and $g(z)$ equal the right-hand side, and consider the circle $|z| = 1 - \delta = C$. On the boundary, $|f(z)| = C$ and

$$|g(z)| \leq \left[\int_0^\infty |e^{-\mu[1-Cexp(i\theta)t]}\, dB(t)\right]^{1/K}$$

$$= \left[\int_0^\infty |e^{-\mu t(1-C\cos\theta - iC\sin\theta}|\, dB(t)\right]^{1/K}$$

$$= \left[\int_0^\infty e^{-\mu(1-C\cos\theta)t}\, dB(t)\right]^{1/K}$$

since $|exp(-iC\sin\theta)| = 1$. Because $\cos\theta \leq 1$, it then follows that

$$|g(z)| \leq \left[\int_0^\infty e^{-\mu(1-C)t}\, dB(t)\right]^{1/K} = I.$$

We now show that $|f(z)| > I$, for then

$$|f(z)| > I \geq |g(z)|.$$

This is equivalent to showing that

$$F(C) = C - \left[\int_0^\infty e^{-\mu(1-C)t}\, dB(t)\right]^{1/K} > 0.$$

But $F(1) = 0$ and $F'(1) = 1 - \mu/K\lambda < 0$. Thus

$$\lim_{C \to 1^-} F(C) > 0 \text{ and } F(1-\delta) > 0.$$

Hence $f(z) + g(z) = 0$ has the same number of roots as $f(z) = 0$ inside $|z| = 1$, namely, 1. □□

We know that the B-matrix model has K roots inside the unit circle, and as promised, we now show that these values are indeed unique.

Theorem 1 *The roots of the characteristic equation of the B-matrix model are unique, with one real root in (0,1) for all values of K and a second real one in (-1,0) only when K is even.*

Proof: Use the form $z^K = \beta(z)$. Then by a geometric argument essentially the same as that for the $G/M/1$ queue used in Figure 5.1 of Gross and Harris (1985), it follows that there exists a unique *real root in* (0,1) for *all* K when

$$\beta'(1) > [d(z^K)/dz]_{z=1}.$$

But this is equivalent to

$$\frac{\mu}{\lambda} > K \qquad \text{or} \qquad \frac{K\lambda}{\mu} < 1,$$

which is true from ergodicity.

For K even, we see that there is an additional real root in (-1,0) with a smaller modulus than the positive root, since z^K is a symmetric function and $0 < \beta(-z) < \beta(z)$ for $z \in (0,1)$.

From Lemma 1, we know that (5) has a unique root inside the unit circle for each $n = 1, ..., K$; call it (r_n, θ_n). But it is also true that (8) has a unique (possibly non-integer) value, n_i, for each pair (r_i, θ_i). Thus if we assume for $i \neq j$ that $(r_i, \theta_i) = (r_j, \theta_j)$, it follows that $n_i = n_j$. But this contradicts the uniqueness of (r, θ) for each n. Therefore all K roots (r_n, θ_n) must be distinct. □□

Remember now that the A-matrix models have characteristic equation

$$z^K = A^*(\lambda - \lambda z) = \alpha(z). \tag{10}$$

As noted earlier, this equation has exactly K roots inside or on $|z| = 1$. If A^* is a rational function whose denominator is of degree r, then it clearly follows that there are $r + K$ roots, with r of them outside the unit circle. For the queueing version of this problem (that is, the model $M/M^{(K)}/1$), $r = 1$ and there is therefore exactly one root outside the unit circle, which is, in fact, real. It turns out that all such A-matrix problems have exactly one root greater than 1 (and therefore have at least one root in the complex domain outside the unit circle). The proof of this follows.

Theorem 2 *The characteristic equation of the A-matrix problem always has one real root greater than 1. Furthermore, there is always an additional real root in* $(-1, 0)$ *for K even, but not fo. K odd.*

Proof: The verification that there is always a real root in $(1, \infty)$ comes from an analysis virtually identical to that used for the B-matrix problem. Similar to before, use the probability generating function $\alpha(z)$ as the right-hand side. Then the first intersection of z^K with this pgf occurs at $z = 1$ whenever $\alpha'(1) > 1$. In addition, there is a guaranteed later intersection since the slope of $\alpha(z)$ must eventually exceed that of z^K at some point.

The existence of a second real root in $(-1, 0)$ for K even follows from the fact that the function $\alpha(z)$ is monotone increasing and z^K is symmetric. Thus the intersection must occur at a point with absolute value less than that of the first intersection on the positive real line, namely, $z = 1$. □□

3 CONCLUDING REMARKS

Of course, we recognize that the location of roots in the complex plane is only part of the problem. It remains to develop effective numerical procedures to take advantage of the information. Our current research is partly devoted to this development.

4 ACKNOWLEDGMENTS

The author wishes to express his sincere thanks to Jahong Xie for her help with some of the proofs contained in this paper.

5 REFERENCES

Abolnikov, L. and Dukhovny, A. (1987). "Necessary and Sufficient Conditions for the Ergodicity of Markov Chains with Transition $\Delta_{m,n}(\Delta'_{m,n})$-Matrix," Jl. Appl. Math. & Simul. 1 13-24.

Bailey, Norman T.J. (1954). "On Queueing Processes with Bulk Service," Jl. Royal Statist. Soc. Ser. B 16 80-87.

Chaudhry, Mohan L. and Templeton, James G.C. (1983). A First Course in Bulk Queues, John Wiley, New York.

Gross, Donald and Harris, Carl M. (1985). Fundamentals of Queueing Theory, John Wiley, New York.

Powell, Warren B. (1985). "Analysis of Vehicle Holding and Cancellation Strategies in Bulk Arrivals, Bulk Service Queues," Transportation Science 19 352-377.

DISTRIBUTION LIST

Copy No.	
1	Office of Naval Research 800 North Quincy Street Arlington, VA 22217 Attention: Scientific Officer, Statistics and Probability Mathematical Sciences Division
2	ONR Resident Representative Joseph Henry Building, Room 623 2100 Pennsylvania Avenue, N.W. Washington, D.C. 20037
3 - 8	Director, Naval Research Laboratory Washington, D.C. 20375 Attention: Code 2627
9 - 20	Defense Technical Information Center Building 5, Cameron Station Alexandria, VA 22314
21 - 29	Author
30	GMU Office of Research